13th International Learning Analytics and Knowledge Conference

Arlington, Texas, USA
March 13-17, 2023

Impact of Non-Cognitive Interventions on Student Learning Behaviors and Outcomes: An analysis of seven large-scale experimental inventions

Kirk P. Vanacore
Worcester Polytechnic Institute
Worcester, Massachusetts, USA
kpvanacore@wpi.edu

Ashish Gurung
Worcester Polytechnic Institute
Worcester, Massachusetts, USA
agurung@wpi.edu

Andrew A. McReynolds
Worcester Polytechnic Institute
Worcester, Massachusetts, USA
aamcreynolds@wpi.edu

Allison Liu
Worcester Polytechnic Institute
Worcester, Massachusetts, USA
aliu2@wpi.edu

Stacy T. Shaw
Worcester Polytechnic Institute
Worcester, Massachusetts, USA
sshaw@wpi.edu

Neil T. Heffernan
Worcester Polytechnic Institute
Worcester, Massachusetts, USA
nth@wpi.edu

ABSTRACT

As evidence grows supporting the importance of non-cognitive factors in learning, computer-assisted learning platforms increasingly incorporate non-academic interventions to influence student learning and learning related-behaviors. Non-cognitive interventions often attempt to influence students' mindset, motivation, or metacognitive reflection to impact learning behaviors and outcomes. In the current paper, we analyze data from five experiments, involving seven treatment conditions embedded in mastery-based learning activities hosted on a computer-assisted learning platform focused on middle school mathematics. Each treatment condition embodied a specific non-cognitive theoretical perspective. Over seven school years, 20,472 students participated in the experiments. We estimated the effects of each treatment condition on students' response time, hint usage, likelihood of mastering knowledge components, learning efficiency, and post-tests performance. Our analyses reveal a mix of both positive and negative treatment effects on student learning behaviors and performance. Few interventions impacted learning as assessed by the post-tests. These findings highlight the difficulty in positively influencing student learning behaviors and outcomes using non-cognitive interventions.

CCS CONCEPTS

• **Applied computing** → **Computer-assisted instruction**; • **Human-centered computing** → *Human computer interaction (HCI).*

KEYWORDS

Computer Assisted Learning Platform, A/B Testing, Causal Inference, Non-Cognitive Factors

LAK 2023, March 13–17, 2023, Arlington, TX, USA

ACM ISBN 978-1-4503-9865-7/23/03.
https://doi.org/10.1145/3576050.3576073

ACM Reference Format:
Kirk P. Vanacore, Ashish Gurung, Andrew A. McReynolds, Allison Liu, Stacy T. Shaw, and Neil T. Heffernan. 2023. Impact of Non-Cognitive Interventions on Student Learning Behaviors and Outcomes: An analysis of seven large-scale experimental inventions. In *LAK23: 13th International Learning Analytics and Knowledge Conference (LAK 2023), March 13–17, 2023, Arlington, TX, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3576050.3576073

1 INTRODUCTION

In recent decades, the use of computer-assisted learning platforms (CALPs) as a complement to traditional classroom instruction has increased dramatically [17, 21, 48]. Learning-experience designers often embed both academic and non-academic supports into CALPs to improve students' learning outcomes. Much of the causal research in learning analytics focuses on the effects of academic interventions in CALPs – such as hints, scaffolding, and performance based-feedback [29, 45, 50, 51] – but less attention has been given to non-cognitive interventions that focus on students' mindsets, emotions, and motivation as they engage in learning activities. There is growing appreciation that "non-cognitive" aspects of education, such as students' motivation, persistence, and meta-cognition, are essential parts of the learning process [23, 61]. Non-cognitive factors are broadly defined as a set of skills and traits that are not analytic or intellectual but are academically relevant [56]. Although these factors are predictive of educational outcomes, attempts to implement non-cognitive interventions in various educational settings, from traditional classrooms and massive open online courses, have shown mixed results in influencing students' behaviors and learning outcomes [19, 25, 70]. As learning-experience designers incorporate non-cognitive interventions into CALPs, research on how these interventions impact students' learning behavior and outcomes is essential to optimizing digital educational environments.

One way to influence non-cognitive factors is by embedding small interventions within a learning activity. In non-educational fields, behavioral scientists have found that embedding simple interventions, such as messages or short activities, "nudge" people toward specific desirable behaviors [37, 62]. Similarly, education researchers studying the impact of 'nudges' meant to influence students' mindsets and motivation and thus impact their academic

behaviors and performance have found mixed results [25, 38, 49]. In a review of social-psychological interventions in education [70], Yeager and Walton note that non-cognitive interventions can seem "magical" because of their potentially long-lasting impacts, but they warn against this view. Instead, they state that we must understand these interventions as "powerful but context-dependent tools" for them to be effective at scale. In that light, we seek to understand the potential impact of these tools in the context of CALPs.

Presently, we explore the use of small non-cognitive interventions in ASSISTments focused on middle school mathematics. Specifically, we investigate the impacts of five experimental interventions that use motivational messaging and small metacognitive exercises based on non-cognitive theories – growth mindset, achievement emotions and control-value, social comparison and self-concept, and metacognition – that have implications for learning. We examine how these interventions impact learning by looking at multiple dimensions of students learning processes. Our findings provide evidence that different types of non-cognitive interventions embedded in online learning systems affect students' learning behaviors and outcomes differently. Practically, these findings can also inform the design of non-cognitive supports within online learning systems that can meaningfully impact student engagement and learning.

2 PRIOR WORK/BACKGROUND

A growing number of studies have investigated the relations between non-cognitive factors and learning [28, 56, 61]. The degree to which these factors are associated with learning outcomes vary. For example, affective factors such as depression, feelings of well-being, and generalized anxiety generally show small correlations with academic performance, whereas domain-specific self-efficacy and domain-specific anxiety are highly correlated with academic achievement [61]. Non-cognitive interventions seek to impact learning by influencing non-cognitive factors, which have theoretical downstream effects on their learning behavior and performance. Below, we describe four common educational theories focused on non-cognitive factors and how they purportedly impact learning behaviors and achievement.

2.1 Mindset Theory

Mindset theory focuses on individuals' beliefs about how abilities can be developed [12]. Individuals can hold either "growth" mindsets, in which they believe their talents can be developed through effort, or "fixed" mindsets, in which they believe their talents are innate and inflexible. According to this theory, people with higher growth mindsets should be more adaptive, particularly when facing difficulties, leading to increased academic achievement.

Studies have found that mindset can influence how students view and engage in learning activities. One early study on mindset found that the mindset portrayed by adults when they praise students' motivation and performance on learning tasks [40]. Research suggests students' mindsets are implicated in academic performance, how they address changes, and respond to academic stress [67]. Growth mindsets may also have indirect effects on achievement through motivational factors, such as increasing motivation to learn [5, 53] and grit [44].

However, there has been considerable debate around the generalizability of growth mindset interventions, as studies have found that effects are highly variable depending on individuals and contexts [68]. Sisk and colleagues [58] conducted two meta-analyses: one on the correlation between mindset and academic achievement (273 studies, n = 365,915) and one on the effect of mindset interventions on academic achievement (43 studies, n = 57,155). They found that over half of the effect sizes reviewed in their meta-analyses were not significant, and those that were significant had an average weak effect along with high degrees of heterogeneity. Specifically, students from lower socioeconomic backgrounds or those who were academically at-risk generally benefited, which has also been reflected in other large-scale studies. For example, a growth mindset predicted achievement in a national sample of 10^{th} grade students in Chile and helped to offset the negative effects of poverty on achievement [9]. Similarly, a large-scale randomized controlled trial investigating the effect of a short online mindset intervention on lower-achieving 9^{th} grade students also found positive impacts on students' math GPA and enrollment in advanced math classes [69]. Still, other recent large-scale studies suggest that growth mindset may only be predictive of achievement among wealthy students and not those from less advantaged families [24]. Thus, it is still unclear how different implementations of growth mindsets will generalize when implemented into a large-scale CALP.

A key aspect of growth mindset is how students respond to failure during learning. Students with growth mindsets should view mistakes as learning opportunities, whereas students with fixed mindsets may view the same mistake as an indication of their poor ability. A study investigating undergraduates in a STEM course found that failure perception was a significant factor associated with changes in students' mindsets over the semester [30]. Students with fixed mindsets at the start of the semester were likely to report academic struggles and shift away further from growth mindsets throughout the semester. Alternatively, students who started the semester with growth mindsets and continued to hold higher growth mindsets throughout the semester reported lower levels of academic struggle the during the semester. This suggests that how students perceive difficulty in learning influences whether they struggle academically.

In order to help students access potential benefits of growth mindsets, many interventions provide scripts for teachers that re-frame students' mistakes as opportunities to learn [57]. In educational games, growth mindset interventions have also increased overall gameplay, positive strategies, and persistence after challenges [43]. As such, growth mindset interventions focused on changing students' perceptions of failure while they engage with CALPs may be effective for improving learning behaviors and outcomes.

2.2 Achievement Emotions & Control-Value Theory

Achievement emotions are those tied to achievement activities or outcomes [46]. The specific emotions that individuals feel during or as a result of learning activities can have significant effects on learning and performance [47, 52]. As part of his Control-Value theory, Pekrun *et al.* [47] posited that individuals' achievement emotions are proximally determined by their appraisals of control

(*i.e.*, their perceived influence over actions and outcomes) and value (*i.e.*, their perceived importance of success). The theory also classified 17 achievement emotions based on their valence (positive vs. negative), activation (activating vs. deactivating), and object focus (activity vs. outcome). For example, joy is classified as a positive, activating emotion focused on outcomes, and frustration is classified as a negative, deactivating emotion focused on activities.

Different emotions have also been found to relate differently with student achievement [7] through mechanisms such as consuming cognitive resources [14, 35], promoting certain strategies [10], or supporting interest and motivation to perform tasks. In general, an emotion's impact on learning depends on where it falls within Pekrun's taxonomy. Positive activating emotions are positively correlated with learning and performance, including outcomes of interest, effort invested, self-regulation of learning, grades, and test scores [31, 47]. Meanwhile, negative deactivating emotions can reduce cognitive resources available for tasks or lead to superficial information processing [46]. However, not all negative emotions are necessarily bad for learning outcomes; for example, a negative-activating emotion such as confusion can benefit learning outcomes if students overcome their confusion [13].

Much prior research on achievement emotions has taken place in controlled lab settings; therefore, it is unclear how non-cognitive interventions focused on achievement emotions will generalize in classrooms and at scale. A potential target for non-cognitive interventions is to promote positive, activating emotions such as joy or hope, which are associated with positive learning outcomes. Another option is to teach students to become more aware of their emotions during or after learning activities to help them identify and regulate those emotions. This emotion labeling is considered part of students' emotion knowledge (*i.e.*, the ability to perceive and label emotions in oneself accurately and others), which is considered a critical non-cognitive skill for students to learn. For example, a meta-analysis of 49 studies with students from ages 3-12 found that emotion knowledge of other people's emotions was correlated with academic performance with an effect size of .32, with stronger associations among middle-class children [65]. However, few studies have investigated how promoting emotion-labeling behaviors can impact their learning outcomes and behaviors.

2.3 Social Comparison Theory & Self-Concept

Social comparison describes any processes through which individuals relate their abilities to others [11]. People engage in social comparison for various reasons, including self-evaluation, self-improvement (to improve their skills), or self-enhancement (to protect or improve their self-esteem). In general, people compare themselves with people who are superior to them in some way, with such comparisons resulting in worsened moods and lower ability appraisals. In contrast, comparisons with people who are lower in ability result in more positive outcomes but are rarer [16]. This tendency has been leveraged by behavioral scientists to influence various behaviors, such as reducing energy conception [2, 41], health-related decisions [32, 37], and tax compliance [3]. These interventions utilize normative data on groups, to nudge individuals towards or away from certain behaviors.

In terms of academic achievement, students' academic self-concepts – perceptions of one's academic abilities – are partly determined by social comparisons of their achievements against their peers [33]. In turn, academic self-concepts have been related to academic achievement [39, 54]. In a classroom context, students can explicitly compare their achievements against individual students while also implicitly measuring themselves against the perceived average ability and performance of their entire school, grade, or classroom [59]. Ability grouping or tracking within schools can further complicate which comparison groups are most salient for students. Indeed, students' social comparisons at the school- or classroom-level have significant effects on students' self-concept. Average school or classroom ability can affect students' individual student's self-concepts even when individual achievement is controlled [34, 55]. Thus, interventions that manipulate the target of social comparison and students' standing relative to a comparison group may be one way to bolster students' academic self-concepts and improve subsequent learning outcomes.

2.4 Metacognition

Metacognition encompasses individuals' knowledge and regulation of their cognition [15, 64]. In terms of academic achievement, having higher metacognitive knowledge can help students understand the factors that affect their academic outcomes and to plan, monitor, and evaluate their learning. Supporting metacognitive strategies during learning activities (*e.g.*, reflecting on one's knowledge prior to or after a learning activity) can help students to think about their learning process [4]. Several prior studies have shown that greater use of metacognitive strategies is positively and strongly associated with higher academic achievement [6, 42, 63, 66]. For example, in the 2009 PISA dataset of 15-year-old students across 65 countries, metacognitive strategies significantly predicted academic achievement when controlling for SES and gender [6].

Confidence judgments are one metacognitive strategy that can help students assess their state of knowledge [36]. Judgments conducted before learning reflect students' perceptions of their current knowledge and ease of learning, which are important components of students' metacognitive self-monitoring. Such judgments require students to think about the requisite knowledge for the task, their own abilities, and the steps they need to take for requisite knowledge and their abilities to align. Teaching students to self-monitor and make knowledge judgments can increase their overall test confidence; however, there is a risk that it can also make them overconfident in their abilities [20]. Thus, non-cognitive interventions that encourage students to judge their confidence while providing timely feedback during problem-solving may be another method for improving their metacognition, which should subsequently ease their learning processes and improve performance.

3 CURRENT STUDY

In the current study, we evaluate the impact of non-cognitive interventions embedded into mastery-learning activities on student learning and learning-related behaviors. The study includes five experiments conducted through ASSISTments. This ASSISTments

platform allows researchers to embed experiments into mastery-based learning activities and problem sets. To date, over 80 experiments have been run through the system. To determine which experiments to include in our analyses, we reviewed all experiments (n = 14) which did not include manipulations with academic features (hints, feedback, problem type, *etc.*). Of these experiments, five were included in the current study because they were conducted with a non-cognitive theoretical basis (*e.g.*, Growth Mindset, social comparison, *etc.*) and included a learning outcome to be used as one of the dependent measures. Each of the selected experiments were conducted in a mastery-based learning activity focused on a specific knowledge component, in which students completed problems until they mastered that knowledge component [60]. This selection process was conducted before extracting the data from the ASSISTments database. Prior to analyzing the data, we preregistered our analyses through the Open Science Foundation (OSF Link). Data and code for the analyse can be found on GitHub (GitHub Link).

3.1 Research Questions

As we are interested in understanding whether non-cognitive interventions affect learning-related behaviors as well as learning measures as outcomes, our research questions focus on a variety of different variables as outcomes:

RQ 1 **Did each intervention increase students' initial response time when solving a problem?**
As response time (defined as the time between viewing the problem and either submitting a response or requesting a hint) is positively correlated with performance [8, 18, 26], we analyzed whether each intervention impacted the amount of time between viewing the problem and taking an action (response time) as a learning-related outcome. This allows us to evaluate whether the interventions cause students to slow down and consider the problem prior to acting, an action that is related to learning and performance [27].

RQ 2 **Did each intervention increase the likelihood of students engaging in hint usage during the activity?**
Providing access to hints has a positive effect on student performance [45, 50], so the likelihood that students utilized hints as a help-seeking behavior was included as an outcome. This allows us to measure the extent to which the interventions cause students to seek assistance as they worked through the activity.

RQ 3 **Did each intervention increase the likelihood that students completed the activity by mastering the knowledge component?**
The goal of each activity in ASSISTments is to master the knowledge components, so we evaluate each intervention's impact on the likelihood that the students reach this goal. Notably, mastery can be viewed as a product of student knowledge entering the activity, learning during the activity, and their willingness to persist through the activity.

RQ 4 **Did each intervention impact student learning as measured by their efficiency in mastering the knowledge component of the activity and performance on a post-test?**
Students required different numbers of problems before reaching mastery and may have experienced different levels of learning during the activity. Therefore, we examined whether the interventions impacted the efficiency in which students learned by examining the difference in the number of problems to mastery, and we evaluated how much learning they experienced during the activity by assessing differences in their performance on a post-test.

4 METHOD

4.1 Experimental Interventions

The five experimental interventions were conducted over eight problem sets. Two of the experiments had multiple treatment conditions, producing a total of seven experimental treatments. None of the interventions overlapped; only one experiment was conducted in each problem set, so students could only be in one intervention at a time. All experiments included a business-as-usual control condition, in which the students worked on the mastery-based learning activities and problem sets without any non-cognitive intervention.

The activities focused on different mathematical skills commonly covered as part of the United States middle school curricula (grades 7-8). These skills include adding decimals, adding and subtracting fractions, percentages, geometry, probability, permutations, and combinations.

Table 1 provides descriptions of each experiment, a label of their theoretical basis, indications of when the interventions were administered during the activity , and the number of problem sets in which the experiment was conducted. In the *embracing mistakes* treatment conditions, students received either *image* or *video* messages encouraging them to adopt a growth mindset by reappraising their mistakes as part of the learning process. The *inspirational quotes* intervention encouraged students to adopt joyful or hopeful emotions as they progressed through the activity by providing motivational messages and positive quotes from celebrities (Albert Einstein, Michael Phelps, and Nicki Minaj). The *social comparison* included two treatment conditions which presented students with normative information about the number of problems that students completed before mastery (*performance*) or the percentage of students who used hints during the activity (*hint usage*). The *emotion labeling* intervention engaged students in a metacognitive reflection on their emotions after they completed the first two problems during the activity. Similarly, the *confidence assessment* intervention encouraged students toward metacognitive reflection, but instead had students reflect on their confidence in solving prior to completing each problem in the activity. Notably, this final intervention was conducted in an experiment that did not include a post-test.

4.2 Data

The data included in our analyses were collected during seven school years in the United States (October 2015 through September 2022). During this time, the assignments were made available to teachers in middle school who used ASSISTments as an instructional tool and assigned these activities to their students as part of their lesson plans. During the experiment, 20,472 students worked

Table 1: Intervention descriptions and theoretical basis

Treatment Conditions	Theory	Description	Administration [1] *Before*	*During*	Number of Problem Sets
Embracing Mistakes					
Image	Growth Mindset	Students were exposed to an image that said "Keep Calm and Learn From Your Mistakes" and a written message that encouraged students to reappraise their mistakes as opportunities to learn prior to starting the mastery learning activity	Yes	No	2
Video	Growth Mindset	Students were exposed to a video that encourages students to reappraise their mistakes as opportunities to learn prior to starting the mastery learning activity	Yes	No	2
Inspirational Quotes	Achievement Emotions & Control-Value Theory	Students were provided with positive messages and inspirational quotes from famous people after they submitted answers answers during the mastery learning activity	No	Yes	2
Social Comparison					
Performance	Social Comparison	Students were told average number of problems completed by their peers to master the content prior to starting the mastery learning activity	Yes	No	1
hint usage	Social Comparison	Students were told the percentage of their peers who used hints in the problem set prior to starting the mastery learning activity	Yes	No	1
Emotion Labeling	Achievement Emotions & Metacognition	Students are asked to generally evaluate their mood using a multiple choice question: "How are you feeling right now? happy; frustrated; relieved; still confused"	No	Yes	1
Confidence Judgements	Metacognition	Students are asked to evaluate their confidence in solving problems upon seeing a problem, but before they are able to submit a response. This occurs three times during the learning activity on three separate problems	No	Yes	2

on the experimental problem sets. Of the total sample, 3,722 students participated in multiple experiments. As students were randomized prior to participation in each experiment, this overlap does not bias the estimates, so all students were included in the analyses. These students participated in a total of 25,220 experimental mastery-based learning activities. The samples for each experiment are shown in Table 2.

Table 2: Experimental Sample Sizes

	Treatment n	Treatment %	Control n	Control %
Embracing Mistakes	2649	66.16%	1355	33.84%
Image	*1341*	*33.49%*		
Video	*1308*	*32.67%*		
Inspirational Quotes	2935	44.54%	3655	55.46%
Social Comparison	544	67.58%	261	32.42%
Performance	*268*	*33.29%*		
Hint Usage	*276*	*34.29%*		
Emotion Labeling	3887	54.02%	3309	45.98%
Confidence Judgement	3306	49.90%	3301	50.10%

4.3 Analytic Approach

To understand the various ways in which each experiment could impact student learning and learning-related behaviors, we utilized five different outcome variables, which align with our research questions. Each variable provides a different perspective on student learning and the student learning process.

[1] Indicates whether the intervention was administered before the mastery-based learning activity, or while the students were completing problems in the activity.

Effect sizes were estimated for each outcome using a series of regression models. We estimated a model for each experiment. Each model compared the treatment condition to the specific control condition associated with that experiment. For experiments with multiple conditions, one model was run for the overall treatment effect and then two subsequent models were estimated for each individual treatment effect. To account for potential inflation of Type I errors due to multiple comparisons, we adjusted the p-values of the effect sizes for family-wise error across the outcomes using the Bonferroni-Holm procedure within each experiment [1].

Equation 1 is the basis used to estimate the treatment effects for each experiment on each outcome. Treatment_i is a binary indicator for whether the student received the treatment. β_1 is the effect of the treatment on the outcome. Specifics of the models vary by each outcome and the details of each outcome and model are described below.

$$\text{outcome}_i = \beta_0 + \beta_1 \text{Treatment}_i + \epsilon_i \tag{1}$$

4.3.1 RQ1 – Impact on Response Time. Response time is an important indicator of learning mathematics as pausing before submitting responses is associated with higher learning outcomes [8, 18]. In theory, pausing allows students to participate in metacognitive reflection, which can help them consider what strategies to apply to the problem [26]. In ASSISTments, response time is the time between entering the problem and the first action, which includes submitting an answer or requesting a hint.

To evaluate the impact of each experimental condition on response time, we estimated linear regressions for each treatment condition. Response time was averaged across all the problems the student completed within each mastery-based learning activity. As response time does not have a normal distribution (skew = 65.46, kurtosis = 6058.54), we used the log response time in the model.

4.3.2 RQ2 – Impact on Hint Usage. ASSISTments allows students to request hints that are developed specifically for each problem in the mastery-based learning activity. The problems in the master learning activity have up to six hints, each providing incrementally more information about how to solve the problem. If students select the final hint, they receive an entire worked example of the problem and the answer. Access to hints increases the probability that students will get the next problem correct [45, 50].

To evaluate the impact of each experimental condition on students' hint usage, we estimated logistic regressions for each treatment condition. We regressed the treatment indicators on whether the student requested and received at least one hint during the mastery-based learning activity. Accessing hints was treated as a binary indicator because of the non-normal distribution of the hints usage (skew = 5.05, kurtosis = 45.12) and only 32.48% of the students across all experiments utilized hints during the activities.

4.3.3 RQ3 – Impact on Knowledge Component Mastery. For each mastery-based learning activity, students needed to reach a threshold of problems correct in a row to advance to the post-test. Thresholds ranged from 3 to 5 problems based on the experiment but were constant across conditions within each experiment. Completing the appropriate number of problems correct in a row indicates that students have mastered the knowledge component for the activity [22]. To assess the effects of each treatment condition on the likelihood that students will master the knowledge component of the activity, we employed logistic regressions.

4.3.4 RQ4 – Impact on Learning Performance. We used two outcomes to evaluate the treatment effects on learning performance: efficiency in mastery and post-test performance. Only students who mastered the knowledge component during the activity had efficiency metrics and were able to participate in the post-test. Students who did not master the knowledge component are considered part of the attrition group. The models for RQ3 serve as tests of attrition balance across conditions. Therefore, if the models from RQ3 show significant differences in the likelihood of mastery between treatment groups, the estimates of the RQ4 models predicting efficiency in mastery and post-test performance may be biased. Even if there is no evidence of attrition imbalance, the estimated effect sizes on efficiency in mastery and post-test performance are still limited to students who mastered the materials.

First, we evaluated how efficient students were in mastering activities' knowledge components. We used the number of problems each student took to master the knowledge component to create an efficiency variable. To ease interpretation, we standardized the number of problems to mastery and multiplied the variable by -1. Therefore, a positive value on the efficiency measure indicates fewer problems to mastery than a negative value, signaling greater efficiency. To evaluate the impact of each experimental condition on students' efficiency, we regressed the treatment indicators on the number of problems students completed to reach mastery.

Second, all but one of the experiments (*confidence judgments*) included post-test problems. The post-tests were brief, including only two or three items. The brevity of the post-test was due to implementation constraints of ecologically valid large-scale experiments, in which long post-tests are not feasible and would likely result in greater attrition. Notably, this is not an ideal evaluation of individual students' learning, but provides and indication of the interventions impact on learning across the entire population, in terms of differences in probability of answering a post-test problem correctly. The post-test problems were designed to be more difficult than mastery learning activities, but to utilize the same skills mastered within the learning activity. The post-test questions required students to transfer the skill learned during the assignment to complex problems often involving multi-step word problems. Each item is intended to assess how much learning occurred in the master-based learning activity.

To evaluate the impact of the treatments on the post-test, we estimated logistic regressions. Because aggregating a limited number of items does not produce a normal distribution, we regressed the treatment indicator on whether the student got each post-test problem correct. Equation 2 delineates this model. To account for variances in student ability, we included a random intercept for each student μ_i. To account for variations in post-test problem difficulty, we included a random intercept for each post-test problem μ_j. The post-test treatment effect is γ_{10}, which is the average difference in the log-odds of students in the treatment group answering a post-test problem correctly relative to students in the control group.

$$\begin{aligned} logit(&\text{Student } i \text{ Gets Post-Test Problem } j \text{ Correct}) \\ &= \gamma_{00} + \gamma_{10}\text{intervention}_i + \mu_i + \mu_j \end{aligned} \tag{2}$$

5 RESULTS

Results from all of the models are displayed in Figure 1, which shows the effect size of each experiment on every outcome, including, confidence intervals, and statistical significance for each experiment on every outcome. The statistical significance indicated in the figure is based on the p-value corrected for family-wise error within each experiment. Notably, few of the interventions impacted students' learning or learning-related behaviors as measured by the studies' outcomes. The results for each experiment's impact are detailed in the sections below.

5.1 Embracing Mistakes Intervention

The embracing mistakes intervention did not significantly impact any of the outcomes. This finding was true regardless of whether the methods of message delivery (image or video) were evaluated individually or as one treatment. Overall, we found no evidence that messaging encouraging students to embrace their mistakes impacts their learning or learning-related behaviors.

5.2 Inspirational Quotes Intervention

Although the inspirational quotes intervention had no statistically significant effect on response time, hint usage, or post-test performance, there were interesting patterns in the treatment's effects on mastery rates and efficiency. Students who were exposed to inspirational quotes were more likely to master the activity's knowledge component (β_1 = 1.13, SE = 0.06, p < 0.001). This effect size is substantial; students in the treatment condition had a .85 probability of mastering the knowledge component compared with .65 for the control group.

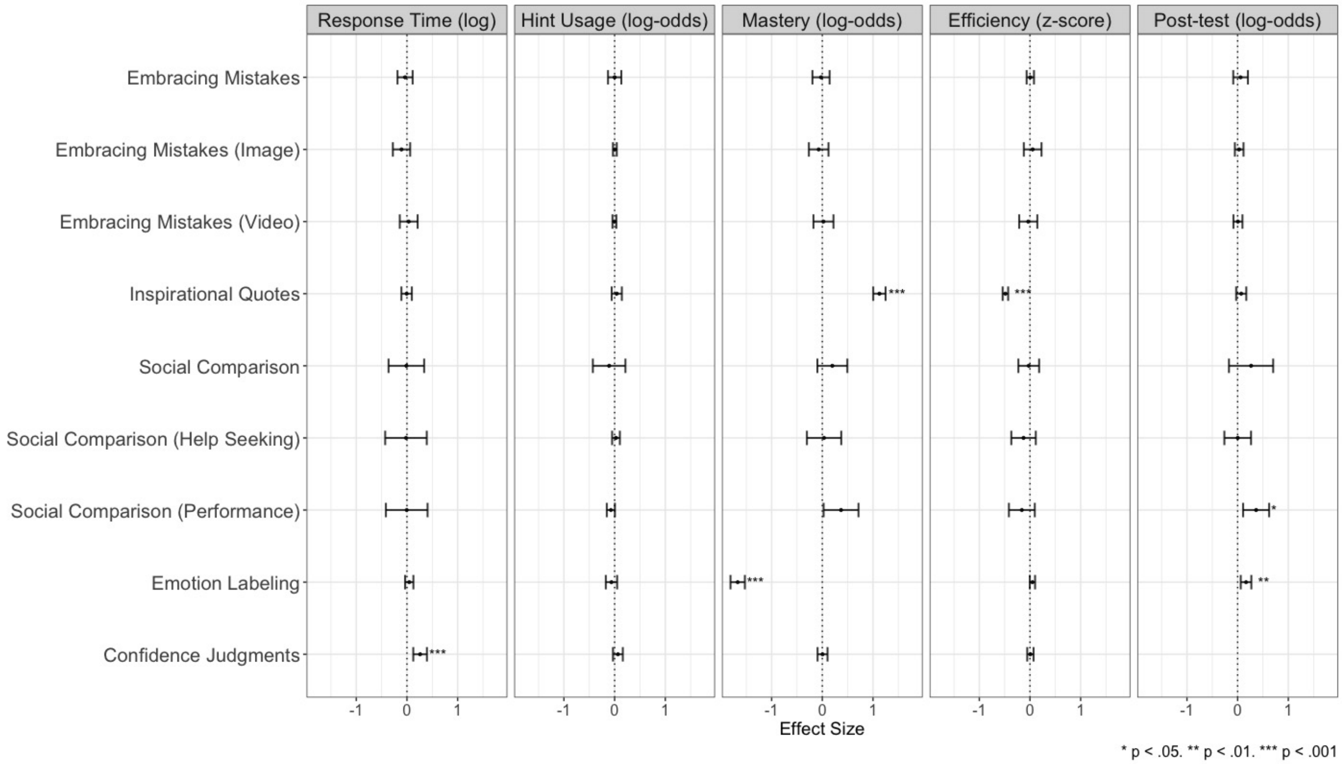

Figure 1: This plot provides the effect size, standard deviation, and statistical significance for each of the experimental conditions on each outcome. Statistical significance is based on p-values adjusted using the Bonferroni-Holm procedure within each experiment.

Of the students who did master the knowledge component, students in the treatment group were significantly less efficient than those in the control group (β_1 = -0.49, SE = 0.03, p = 0.001). Notably, the decrease in efficacy in the treatment group may be related to an increase in the number of students mastering. If the treatment inspired students who would otherwise give up to persist and complete more problems, ultimately mastering the knowledge component, this would drive up the average number of problems completed for the treatment group, thus explaining the decrease in efficiency.

Overall the treatment did not have a significant effect on student learning as measured by the post-test. However, attrition imbalance may have biased this result. Since only the students who mastered the activities' knowledge component received the post-test, and the treatment encouraged students who would not have mastered to do so, this result may be biased towards the control group.

5.3 Social Comparison Intervention

Overall, there were no significant differences in learning or learning-related behaviors between the students who received the social comparison treatments and those in the control group. However, when the social comparison treatments were compared to the control individually, an interesting pattern emerged; the *performance* messaging seemed to have some impact on performance, whereas the hint usage messaging had no detectable impact on *hint usage*.

Students who received *performance* messaging outperformed the control on the post-test (β_1 = 0.36, SE = 0.13, p = 0.022). Notably, the *performance* messaging focused on efficiency by presenting the number of problems the average student completed before reaching mastery did not significantly affect students' efficiency. Yet, this condition had only a marginally non-significant effect on mastery; the effect size was moderate (β_1 = 0.37, SE = 0.17, p =0.141), and this effect was non-significant only after the family-wise error correction and the majority of the confidence interval lies in the positive direction. This result suggests that fine-tuning the messaging may produce a better effect on mastery in future interventions.

The *hint usage* condition had a non-significant effect on learning or learning-related behaviors. Interestingly, this condition provided students with a message that the majority of students utilized hints, yet it had a non-significant effect on (β_1 = 0.027, SE = 0.04, p > 0.999).

5.4 Emotion Labeling Intervention

Emotion labeling had statistically significant effects on mastery, and the post-test but not on response time, hint, usage and efficiency. The effect on mastery was negative and substantial (β_1 = -1.67, SE = 0.07, p < 0.001). The probability of mastery for the treatment condition was .68, whereas the probability for the control was .92.

The effect on the post-test was significant and positive (β_1 = 0.17, SE = 0.05, p = 0.006). However, the effect may be caused by the attrition imbalance as fewer students in the treatment group mastered the knowledge component and had the opportunity to take the post-test. Therefore the causal effect of the treatment on the post-test is questionable. The pool of students who mastered the knowledge component may have been generally higher performing in the treatment than the control, thus biasing the results.

5.5 Confidence Judgments Intervention

The confidence judgments invention did not significantly affect hint usage, mastery, or efficiency. The impact on response time was (β_1 = 0.34, SE = 0.07, p < 0.001). This suggests that the treatment condition may have slowed students' response time by encouraging them to think more about the problems. Yet, this increase in response time did not produce changes in other learning-related behaviors as measured by hint usage, mastery, or efficiency. The experiment did not include a post-test.

6 DISCUSSION

The non-cognitive interventions analyzed in this paper produced mixed results regarding their impacts on learning and learning-related behaviors. Notably, none of the interventions produced consistently positive significant results across all the outcomes. The *social comparison* intervention, which provided normative messaging about student performance, was the only intervention to produce a positive result on the post-test while also passing the threshold of attrition imbalance by not having differences in mastery rates. Emotion labeling had negative impacts on both mastery rates and efficiency. Although students in this condition outperformed the control on the post-test, the reduced mastery rates in the treatment potentially biased this result, thus confounding the causal inference. Furthermore, *inspirational quotes* increased students' likelihood of mastery but failed to produce a positive effect on learning as measured by the post-test. Nevertheless, this may still be a favorable finding as students in treatment who would stop before mastering had they been in control may still have performed as well as control students. In sum, these experiments illustrated the difficulty of constructing non-cognitive interventions that positively impact students' behavior and performance on learning activities.

This difficulty in constructing impactful non-cognitive interventions may be due to the nature of non-cognitive factors. For example, for students to adopt a growth mindset while participating in a learning activity, they may need more extensive changes to their learning culture that transcend a small message at the beginning of an activity. One large-scale online mindset intervention found that the intervention's effects on math performance were dependent, in part, on whether the student's peer groups were supportive of a growth mindset [69]. Hence, a student who entered the *embracing mistakes* experiment may have already adopted a mindset based on their social learning environment and prior experiences that influences their perceptions of mistakes more than their exposure to the experimental messaging. Therefore, to impact student learning, non-cognitive intervention may require more robust changes to students learning environments than small messages and activities embedded in a CALP. Future work should examine whether CALPs can broadly influence learning environments to affect non-cognitive factors and, subsequently, learning behaviors and outcomes.

One surprising finding was the nonsignificant effects of the *confidence judgments* intervention. This intervention was intended to inspire a metacognitive reflection on the alignment between students' abilities and problem difficulty. In theory, this should help students assess their knowledge state, which should ease the learning process [36]. In practice, this reflection did not affect the likelihood that students would master the knowledge component or improve their efficiency in learning. Notably, the other metacognitive intervention evaluated in this study, *emotion labeling*, had a negative impact on mastery and efficiency. These interventions may have been demotivating for students who either had low confidence in their ability through the activity or started with high confidence but had difficulty mastering the concept. There is evidence that metacognitive reflection could lead students to overconfidence [20]. The effects may differ based on students' levels of confidence or emotional state. Future analyses should focus on understanding the behaviors of these students based on their responses to the confidence judgment and emotion identification questions. Furthermore, qualitative data, such as information produced by 'think-alouds' and interviews, may help understand why emotion labeling can produce adverse effects.

Although our study provides interesting causal data on the impact of non-cognitive interventions in CALPs, there are some notable limitations. First, the lack of a multi-item post-test reduces the overall accuracy of assessing student learning and should be interpreted cautiously. It is infeasible to provide long muti-item post-test assessments in large-scale ecologically valid experiments, especially to estimate the effects of minor interventions embedded in short assignments. Future work should address this difficulty by developing valid and feasibly implemented measures in these contexts. Second, the implicit attrition caused by the master-learning activity potentially biases the results. Since only the students who master the activity get to take the post-test, our inferences about the impact of the intervention on learning as measured by the post-test is limited to a subset of the student population. Furthermore, when the intervention impacts the likelihood that students master the activity, this biases any estimates of effects on learning. Future work should consider using quasi-experimental methods to account for these imbalances or providing post-tests to all students regardless of mastery, perhaps at a set point in the activity as opposed to after mastery. Finally, our ability to understand the mechanisms of these interventions or make inferences about who may benefit from these interventions is limited by our lack of knowledge about the students who use many CALPs. More demographic information should be collected to understand the potential differential impact across populations.

7 CONCLUSION

Our work highlights the difficulty of developing non-cognitive interventions which positively impact students' academic behaviors and outcomes. Our findings confirm Yeager and Walton's assertion that non-cognitive interventions are not "magical" solutions but tools

that only work in specific contexts with specific implementations [70]. Learning experience designers should note that applying non-cognitive theories while developing theoretically helpful features for students may result in adverse effects in practice. For example, we found that the emotional labeling intervention negatively affected students' likelihood of mastering knowledge components; hence, embedding meta-cognitive tasks into CALPS learning activities should be done with caution.

Our findings suggest that minor modifications in activity designs may be insufficient to change students' motivation and mindset in ways that impact behavior and learning. Designers may have to take global approaches to implement non-cognitive theories when building CALPs by considering how the entire program can be infused with motivational, mindset-oriented, and metacognitive content. Furthermore, simply changing the CALPs environment may not produce desired effects if the students' broader learning environments are not influenced (as found in [69]).

Finally, as we found both positive and negative results for these non-cognitive interventions, designers should always test the impact as they develop new features or more global meta-cognitive changes to programs. This process will ensure alignment between theory and practice, resulting in positive outcomes for learners.

ACKNOWLEDGMENTS

We would like to thank Korinn Ostrow, NSF (*e.g.*, 2118725, 2118904, 1950683, 1917808, 1931523, 1940236, 1917713, 1903304, 1822830, 1759229, 1724889, 1636782, & 1535428), IES (*e.g.*, R305N210049, R305D210031, R305A170137, R305A170243, R305A180401, & R305A120125), GAANN (*e.g.*, P200A180088 & P200A150306), EIR (U411B190024 & S411B210024), ONR (N00014-18-1-2768), and Schmidt Futures. None of the opinions expressed here are that of the funders.

REFERENCES

[1] Hervé Abdi. 2010. Holm's sequential Bonferroni procedure. *Encyclopedia of research design* 1, 8 (2010), 1–8.
[2] Hunt Allcott and Judd B Kessler. 2019. The welfare effects of nudges: A case study of energy use social comparisons. *American Economic Journal: Applied Economics* 11, 1 (2019), 236–76.
[3] Armenak Antinyan and Zareh Asatryan. 2019. Nudging for tax compliance: A meta-analysis. *ZEW-Centre for European Economic Research Discussion Paper* 19-055 (2019).
[4] Roger H. Bruning, Gregory J. Schraw, and Monica M. Norby. 2011. *Cognitive psychology and instruction.* Pearson.
[5] Jeni L Burnette, Michelle V Russell, Crystal L Hoyt, Kasey Orvidas, and Laura Widman. 2018. An online growth mindset intervention in a sample of rural adolescent girls. *British Journal of Educational Psychology* 88, 3 (2018), 428–445.
[6] Gregory L Callan, Gregory J Marchant, W Holmes Finch, and Rachel L German. 2016. Metacognition, strategies, achievement, and demographics: Relationships across countries. *Educational Sciences: Theory & Practice* 16, 5 (2016).
[7] Jesús Camacho-Morles, Gavin R Slemp, Reinhard Pekrun, Kristina Loderer, Hanchao Hou, and Lindsay G Oades. 2021. Activity achievement emotions and academic performance: A meta-analysis. *Educational Psychology Review* 33, 3 (2021), 1051–1095.
[8] Jenny Yun-Chen Chan, Erin R Ottmar, and Ji-Eun Lee. 2022. Slow down to speed up: Longer pause time before solving problems relates to higher strategy efficiency. *Learning and Individual Differences* 93 (2022), 102–109.
[9] Susana Claro, David Paunesku, and Carol S Dweck. 2016. Growth mindset tempers the effects of poverty on academic achievement. *Proceedings of the National Academy of Sciences* 113, 31 (2016), 8664–8668.
[10] Gerald L Clore and Jeffrey R Huntsinger. 2009. How the object of affect guides its impact. *Emotion Review* 1, 1 (2009), 39–54.
[11] Pieternel Dijkstra, Frederick X Gibbons, and Abraham P Buunk. 2010. Social comparison theory. *Social psychological foundations of clinical psychology* (2010).
[12] Carol S Dweck. 2006. *Mindset: The new psychology of success.* Random House.
[13] Sidney D'Mello, Blair Lehman, Reinhard Pekrun, and Art Graesser. 2014. Confusion can be beneficial for learning. *Learning and Instruction* 29 (2014), 153–170.
[14] Henry C Ellis and Patricia W Ashbrook. 1989. The" state" of mood and memory research: A selective review. *Journal of Social Behavior and Personality* 4, 2 (1989), 1.
[15] John H Flavell. 1979. Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist* 34, 10 (1979), 906.
[16] JP Gerber, Ladd Wheeler, and Jerry Suls. 2018. A social comparison theory meta-analysis 60+ years on. *Psychological bulletin* 144, 2 (2018), 177.
[17] Lucinda Gray and Laurie Lewis. 2021. Use of Educational Technology for Instruction in Public Schools: 2019–20. https://nces.ed.gov/pubs2021/2021017Summary.pdf
[18] Ashish Gurung, Anthony F. Botelho, and Neil T. Heffernan. 2021. Examining Student Effort on Help through Response Time Decomposition. In *LAK21: 11th International Learning Analytics and Knowledge Conference* (Irvine, CA, USA) *(LAK21)*. Association for Computing Machinery, New York, NY, USA, 292—-301. https://doi.org/10.1145/3448139.3448167
[19] Leslie Morrison Gutman and Ingrid Schoon. 2013. The impact of non-cognitive skills on outcomes for young people. A literature review. (2013).
[20] Jessica D Huff and John L Nietfeld. 2009. Using strategy instruction and confidence judgments to improve metacognitive monitoring. *Metacognition and Learning* 4, 2 (2009), 161–176.
[21] Gallop Inc. 2019. Education Technology Use in Schools. https://www.newschools.org/wp-content/uploads/2020/03/Gallup-Ed-Tech-Use-in-Schools-2.pdf
[22] Kim Kelly, Yan Wang, Tamisha Thompson, and Neil Heffernan. 2016. Defining mastery: Knowledge tracing versus n-consecutive correct responses. *Student Modeling From Different Aspects* (2016), 39.
[23] Myint Swe Khine and Shaljan Areepattamannil. 2016. *Non-cognitive skills and factors in educational attainment.* Springer.
[24] Ronnel B King and Jose Eos Trinidad. 2021. Growth mindset predicts achievement only among rich students: examining the interplay between mindset and socioeconomic status. *Social Psychology of Education* 24, 3 (2021), 635–652.
[25] René F Kizilcec, Justin Reich, Michael Yeomans, Christoph Dann, Emma Brunskill, Glenn Lopez, Selen Turkay, Joseph Jay Williams, and Dustin Tingley. 2020. Scaling up behavioral science interventions in online education. *Proceedings of the National Academy of Sciences* 117, 26 (2020), 14900–14905.
[26] Bracha Kramarski and Mary Gutman. 2006. How can self-regulated learning be supported in mathematical E-learning environments? *Journal of Computer Assisted Learning* 22, 1 (2006), 24–33.
[27] Alyssa P Lawson, Arineh Mirinjian, and Ji Y Son. 2019. Can preventing calculations help students learn math? *Journal of Cognitive Education and Psychology* 17, 2 (2019), 178–197.
[28] J Lee and L Stankov. 2013. Higher-order structure of motivation, self-beliefs, learning strategies, and attitudes toward school and its prediction of PISA 2003 mathematics scores. *Learning and Individual Differences* 26 (2013), 119–130.
[29] Jia Li. 2010. Learning vocabulary via computer-assisted scaffolding for text processing. *Computer Assisted Language Learning* 23, 3 (2010), 253–275.
[30] Lisa B Limeri, Jun Choe, Hannah G Harper, Hannah R Martin, Annaleigh Benton, and Erin L Dolan. 2020. Knowledge or abilities? How undergraduates define intelligence. *CBE—Life Sciences Education* 19, 1 (2020), ar5.
[31] Elizabeth A Linnenbrink. 2007. The role of affect in student learning: A multi-dimensional approach to considering the interaction of affect, motivation, and engagement. In *Emotion in education.* Elsevier, 107–124.
[32] Blake M. DiCosola III and Gina Neff. 2022. Nudging Behavior Change: Using In-Group and Out-Group Social Comparisons to Encourage Healthier Choices. In *CHI Conference on Human Factors in Computing Systems.* 1–14.
[33] Herbert W Marsh, Ulrich Trautwein, Oliver Lüdtke, and Olaf Köller. 2008. Social comparison and big-fish-little-pond effects on self-concept and other self-belief constructs: Role of generalized and specific others. *Journal of Educational Psychology* 100, 3 (2008), 510.
[34] Herbert W Marsh, Ulrich Trautwein, Oliver Lüdtke, and Olaf Köller. 2008. Social comparison and big-fish-little-pond effects on self-concept and other self-belief constructs: Role of generalized and specific others. *Journal of Educational Psychology* 100, 3 (2008), 510.
[35] Jörg Meinhardt and Reinhard Pekrun. 2003. Attentional resource allocation to emotional events: An ERP study. *Cognition and Emotion* 17, 3 (2003), 477–500.
[36] Christoph Mengelkamp and Maria Bannert. 2012. *Confidence Judgments in Learning.* Springer US, Boston, MA, 756–759. https://doi.org/10.1007/978-1-4419-1428-6_1726
[37] Katherine L Milkman, Linnea Gandhi, Mitesh S Patel, Heather N Graci, Dena M Gromet, Hung Ho, Joseph S Kay, Timothy W Lee, Jake Rothschild, Jonathan E Bogard, et al. 2022. A 680,000-person megastudy of nudges to encourage vaccination in pharmacies. *Proceedings of the National Academy of Sciences* 119, 6 (2022), e2115126119.
[38] Negar Mohammadhassan, Antonija Mitrovic, and Kourosh Neshatian. 2022. Investigating the effect of nudges for improving comment quality in active video watching. *Computers & Education* 176 (2022), 104340.

[39] Jens Möller, Steffen Zitzmann, Friederike Helm, Nils Machts, and Fabian Wolff. 2020. A meta-analysis of relations between achievement and self-concept. *Review of Educational Research* 90, 3 (2020), 376–419.
[40] Claudia M Mueller and Carol S Dweck. 1998. Praise for intelligence can undermine children's motivation and performance. *Journal of personality and social psychology* 75, 1 (1998), 33.
[41] Erica Myers and Mateus Souza. 2020. Social comparison nudges without monetary incentives: Evidence from home energy reports. *Journal of Environmental Economics and Management* 101 (2020), 102315.
[42] VC Onu, M Eskay, JN Igbo, N Obiyo, and O Agbo. 2012. Effect of Training in Math Metacognitive Strategy on Fractional Achievement of Nigerian Schoolchildren. *Online Submission* (2012).
[43] Eleanor O'Rourke, Kyla Haimovitz, Christy Ballweber, Carol Dweck, and Zoran Popović. 2014. Brain points: A growth mindset incentive structure boosts persistence in an educational game. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 3339–3348.
[44] Daeun Park, Eli Tsukayama, Alisa Yu, and Angela L Duckworth. 2020. The development of grit and growth mindset during adolescence. *Journal of Experimental Child Psychology* 198 (2020), 104889.
[45] Thanaporn Patikorn and Neil T Heffernan. 2020. Effectiveness of crowd-sourcing on-demand assistance from teachers in online learning platforms. In *Proceedings of the Seventh ACM Conference on Learning@ Scale*. 115–124.
[46] Reinhard Pekrun. 2006. The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational psychology review* 18, 4 (2006), 315–341.
[47] Reinhard Pekrun, Thomas Goetz, Wolfram Titz, and Raymond P Perry. 2002. Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational psychologist* 37, 2 (2002), 91–105.
[48] Anthony G Picciano, Jeff Seaman, Peter Shea, and Karen Swan. 2012. Examining the extent and nature of online learning in American K-12 education: The research initiatives of the Alfred P. Sloan Foundation. *The internet and higher education* 15, 2 (2012), 127–135.
[49] Simone Plak, Chris van Klaveren, and Ilja Cornelisz. 2022. Raising student engagement using digital nudges tailored to students' motivation and perceived ability levels. *British Journal of Educational Technology* (2022).
[50] Ethan Prihar, Thanaporn Patikorn, Anthony Botelho, Adam Sales, and Neil Heffernan. 2021. Toward Personalizing Students' Education with Crowdsourced Tutoring. In *Proceedings of the Eighth ACM Conference on Learning@ Scale*. 37–45.
[51] Ethan Prihar, Manaal Syed, Korinn Ostrow, Stacy Shaw, Adam Sales, and Neil Heffernan. 2022. Exploring Common Trends in Online Educational Experiments. In *Proceedings of the 15th International Conference on Educational Data Mining*. 27.
[52] Kathleen M Quinlan. 2016. How emotion matters in four key relationships in teaching and learning in higher education. *College Teaching* 64, 3 (2016), 101–111.
[53] Emily Rhew, Jody S Piro, Pauline Goolkasian, and Patricia Cosentino. 2018. The effects of a growth mindset on self-efficacy and motivation. *Cogent Education* 5, 1 (2018), 1492337.
[54] Carl M Rogers, Monte D Smith, and J Michael Coleman. 1978. Social comparison in the classroom: the relationship between academic achievement and self-concept. *Journal of educational psychology* 70, 1 (1978), 50.
[55] Carl M Rogers, Monte D Smith, and J Michael Coleman. 1978. Social comparison in the classroom: the relationship between academic achievement and self-concept. *Journal of educational psychology* 70, 1 (1978), 50.
[56] Jeffrey A Rosen, Elizabeth J Glennie, Ben W Dalton, Jean M Lennon, and Robert N Bozick. 2010. *Noncognitive skills in the classroom: New perspectives on educational research*. RTI Press.
[57] Heather Savvides and Caroline Bond. 2021. How does growth mindset inform interventions in primary schools? A systematic literature review. *Educational Psychology in Practice* 37, 2 (2021), 134–149.
[58] Victoria F. Sisk, Alexander P. Burgoyne, Jingze Sun, Jennifer L. Butler, and Brooke N. Macnamara. 2018. To What Extent and Under Which Circumstances Are Growth Mind-Sets Important to Academic Achievement? Two Meta-Analyses. *Psychological Science* 29, 4 (2018), 549–571. https://doi.org/10.1177/0956797617739704 arXiv:https://doi.org/10.1177/0956797617739704 PMID: 29505339.
[59] Einar M Skaalvik and Sidsel Skaalvik. 2002. Internal and external frames of reference for academic self-concept. *Educational Psychologist* 37, 4 (2002), 233–244.
[60] Robert E. Slavin. 1987. Mastery Learning Reconsidered. *Review of Educational Research* 57, 2 (Jun 1987), 175–213. https://doi.org/10.3102/00346543057002175
[61] Lazar Stankov and Jihyun Lee. 2014. Quest for the best non-cognitive predictor of academic achievement. *Educational Psychology* 34, 1 (2014), 1–8. https://doi.org/10.1080/01443410.2013.858908
[62] Richard H Thaler and Cass R Sunstein. 2021. *Nudge*. Yale University Press.
[63] Inger Throndsen. 2011. Self-regulated learning of basic arithmetic skills: A longitudinal study. *British Journal of Educational Psychology* 81, 4 (2011), 558–578.
[64] Marcel VJ Veenman, Van Hout-Wolters, HAM Bernadette, and Peter Afflerbach. 2006. Metacognition and learning: Conceptual and methodological considerations. *Metacognition and learning* 1, 1 (2006), 3–14.
[65] Katharina Voltmer and Maria von Salisch. 2017. Three meta-analyses of children's emotion knowledge and their school success. *Learning and Individual Differences* 59 (2017), 107–118.
[66] Anneke Vrugt and Frans J Oort. 2008. Metacognition, achievement goals, study strategies and academic achievement: pathways to achievement. *Metacognition and learning* 3, 2 (2008), 123–146.
[67] David Scott Yeager and Carol S Dweck. 2012. Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational psychologist* 47, 4 (2012), 302–314.
[68] David S Yeager and Carol S Dweck. 2020. What can be learned from growth mindset controversies? *American psychologist* 75, 9 (2020), 1269.
[69] David S Yeager, Paul Hanselman, Gregory M Walton, Jared S Murray, Robert Crosnoe, Chandra Muller, Elizabeth Tipton, Barbara Schneider, Chris S Hulleman, Cintia P Hinojosa, et al. 2019. A national experiment reveals where a growth mindset improves achievement. *Nature* 573, 7774 (2019), 364–369.
[70] David S Yeager and Gregory M Walton. 2011. Social-psychological interventions in education: They're not magic. *Review of educational Research* 81, 2 (2011), 267–301.